MIT LIBRARIES
3 9080 00658182 8

MASSACHUSETTS INSTITUTE of TECHNOLOGY
LIBRARIES

HD28
.M414
0.3168-
90

DEWEY

JUL 26 1990

WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

MONITORING PROGRESS IN R&D COMMUNITIES

by

Michael A. Rappa and Koenraad Debackere

June 1990 WP# 3168-90-BPS

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

MONITORING PROGRESS IN R&D COMMUNITIES

by

Michael A. Rappa and Koenraad Debackere

June 1990 WP# 3168-90-BPS

Michael Rappa is assistant professor of management in the Management of Technology Group, at the Alfred P. Sloan School of Management, Massachusetts Institute of Technology. Koenraad Debackere is a doctoral candidate in the Vlerick School of Management, Gent University. This paper was presented at the International Conference on Science and Technology Indicators, Bielefeld, West Germany, June 10-11, 1990.

M.I.T. LIBRARIES
JUL 2 6 1990.
RECEIVED

ABSTRACT

This paper discusses the use of electronic bibliographic and patent databases as a source of data for monitoring progress in R&D communities. Specifically, the databases are used to access information regarding who the researchers are, what they are working on, who the are working with, where they are working and when. Given this information certain structural and behavioral characteristics of an R&D community are analyzed, including patterns of growth in participation, organizational and geographic distribution, mobility in and out of the field and between organizations, and the collaborative network that evolves among researchers.

A conceptual framework is proposed that can be used for interpreting the rate of progress achieved by a community over time. It is argued that monitoring R&D communities via electronic databases may provide R&D managers with an objective source of information that can aid in them in a firm's R&D investment decisions and in formulating technology strategies.

Monitoring progress in R&D communities

M.A. Rappa and K. Debackere

1. Introduction

The rapid progress of technology continues with little sign of abatement. Over the last several decades there have been astonishing advances across a wide range of technologies, and one can be fairly confident that the next several decades will be equally prolific. Today researchers around the globe are working intently on developing ideas that may create new branches of technological practice and could ultimately transform industry in ways which are hard for most of us to imagine. The ability of managers and policymakers to comprehend the pace and the direction of technological advancement will largely determine a firm's or nation's competitive performance in world markets into the next century. This is no small task. Historical accounts of industrial evolution, such as with the development of semiconductors, videocassette recorders, and personal computers, show the immense difficulties some firms encounter when confronted by new technologies. Undoubtedly, there is an obvious need to enhance our understanding of the way in which new technologies emerge. Different methodologies, ranging from qualitative case studies to extremely sophisticated quantitative forecasting models, have been used to achieve this end. Bibliometric techniques have acquired their proper place in this arena of methodologies.

This paper demonstrates how one can use electronic bibliographic and patent databases as a source of data for monitoring technological progress. Specifically, the databases are used to access information regarding who the researchers are, what they are working on, who they are working with, where they are working and when. The relevant unit of analysis in this type of bibliometric analyses is the R&D community, i.e. the community of scientists and engineers, working on the development of the particular technology under investigation. Using the the literature as a source of data on the R&D community members enables us to analyze certain structural and behavioral characteristics of the community, including patterns of growth in participation, organizational and

geographic distribution, mobility in and out of the field and between organizations, and the collaborative network that evolves among researchers.

The paper starts with a discussion of how one may use the literature in a manner different from volume counting or citation analysis. We then illustrate how the literature data are converted into a number of summary statistics which reflect the longitudinal changes within a R&D community. Next, some of the hurdles of this type of bibliometric methodology, as encountered during our analyses of over a dozen of R&D communities, are discussed. Finally, the discussion section deals with how the methodology can be used to the advantage of the R&D management process.

2. An alternative way of looking at the literature

When, in 1917, Cole and Eaves published their study of the historical development of comparative anatomy, they were among the first to utilize the published literature in order to construct a quantitative profile of the progress of a field of research.[1] Given the technology of the day, it was a truly painstaking effort by the two British scientists, who culled nearly 6,500 books and papers to compile their data. The result was a statistical account of the ebb and flow of research in comparative anatomy over three centuries, since 1550, with disaggregation of the data by country and topical area within the field. The pioneering effort of Cole and Eaves was joined subsequently by two agricultural scientists, Wilson and Fred of the University of Wisconsin, who published a study in 1935 of the growth and development of research on the subject of nitrogen

[1]F.J. Cole and N.B. Eales, "The History of Comparative Anatomy, Part I.--A Statistical Analysis of the Literature," *Science Progress* 11 (1917): 578-96.

fixation by plants.[2]

It was not for another two decades that the study of the scientific literature was to regain serious attention. The one most responsible for this resurgence of interest was Price, who in 1956 used the physics literature to document the "exponential growth of science."[3] Price's work brought legitimacy to the scholarly examination of literature, partly because his conclusions coincided with, and indeed reinforced, the widely held belief that the proliferation of science seemingly knew no bounds. In recent years, studies of emerging fields of science and technology using bibliometric data have become fairly common. The nature of this research is quite diverse and there is to date no single research method that has come to the forefront. However, in general, there are two basic types of studies. The first, built directly on the work of Price and his predecessors, seeks to model the growth of a field by measuring annual publication volume.[4] The second is quite different in that it uses the citation as a unit of analysis.[5] Citation-based studies seek to understand the interlocking nature of citation patterns as a means for observing the growth and the development of clusters of researchers who may ultimately come to form the basis for a new field. Both types of studies cover a wide range of topics in science and technology, although attention to the latter has been increasing. This has led to a more intensive examination of the patent literature, both in terms of the

---

[2]P.W. Wilson and E.B. Fred, "The Growth Curve of a Scientific Literature: Nitrogen Fixation by Plants," *Scientific Monthly* 41 (1935): 240-50.

[3]D.J. Price, "The Exponential Growth Curve of Science," *Discovery* 17 (1956): 240-43.

[4]For example, see A. Granberg, "A Bibliometric Survey of Fiber-Optics Research in Sweden, West-Germany, and Japan," Monograph, Research Policy Institute, University of Lund, Sweden, 1985.

[5]The notion of "bibliographic coupling" was proposed by Kessler in 1962. Since then, numerous co-citation studies have been published, including the pioneering work of Small and his colleagues. For a recent example, see H. Small and E. Greenlee, "Collagen Research in the 1970s," *Scientometrics* 10 (1986): 95-117. A variant of co-citation analysis is co-word analysis (alone or in combination with co-citation) sometimes referred to as "science mapping." See P. Healy, H. Rothman, and P.K. Hoch, "An Experiment in Science Mapping for Research Planning," *Research Policy* 15 (1986) 233-251.

volume of patents and patent citations.[6]

Limitations of bibliometric studies

The studies described in the previous section have led the way in revealing the promise, as well as the limitation, of using bibliometric data for the analysis of emerging fields of science and technology. One fundamental question which arises in this respect is whether bibliometric data can truly reveal anything about *technological* activity, given the belief that it is contrary to the profit-orientation of a firm to publicly disclose the results of its research, and therefore, that most technological activity is not likely to be visible in the literature. However, setting aside the obvious point that bibliometric data can, in fact, include the patent literature, this question is right at the heart of the recently emerging conceptual position that the development of a new technology is not only an organizational activity, though appears to be a communal activity as well.[7] The notion of a R&D community implies a blending of science and technology, of academic and industrial pursuits, and furthermore, suggests a situation where papers and patents are essentially two different mechanisms for staking intellectual claims.[8] Moreover, as far as industrial researchers are concerned, it does not necessarily follow that publishing is an act of complete disclosure, which automatically diminishes the value of information and knowledge to its producer. There is room for

---

[6]For an overview, see B.L. Blasberg, "Patents and the Measurement of Technological Change: A Survey of the Literature," *Research Policy* 16 (1987): 131-41. See also R.M. Wilson, "Patent Analysis using Online Databases: Technological Trend Analysis," *World Patent Information* 9 (1987): 18-26.

[7]Constant was the first to develop the notion that technology development occurs within the context of a broader community of practitioners, see E.W. Constant, *The Origins of the Turbojet Revolution*, (Baltimore: The Johns Hopkins University Press, 1980). This community of practitioners encompasses scientists and engineers working in a diversity of organizations on an interrelated set of scientific and technological problems. This "community of practitioners" has recently been coined with the term "R&D community", see for example Rappa and Debackere, "The Emergence of a New Technology: the Case of Neural Networks," *MIT Working Paper* WP# 3031-89-BPS.

[8]For a compelling treatise on the unification of science and technology, see B. Latour, *Science in Action: How to Follow Scientists and Engineers Through Society* (Cambridge: Harvard University Press, 1987).

both openness and secrecy in the exchange of ideas that occurs between researchers in journals and at conferences. In light of similar arguments, it is not surprising that bibliometric data not only reflect scientific activity, but technological activity as well. However, other criticisms to the use of bibliometric data may be even more threatening.

First of all, publication volume counting and citation patterns analysis alone cannot adequately reflect the multitude of *process-related* factors that may play a significant role in the emergence of a new field. Furthermore, the theoretical underpinning for the use of bibliometric data is not always very clear. The basic idea behind most analyses is that there is a positive relationship between the volume of publication in a field and the rate of progress, and between the quality of the work and the amount of citations. While the logic certainly is appealing, more work is needed to flesh out a theory of emerging fields and the connection between theoretical constructs and the empirical measures derived from bibliometric data.

Although these criticisms have some truth to them, it would be unfortunate if they were to dampen interest in the use of bibliometric data for the study of emerging technologies. The literature of a field is a rich source of information about new technologies, not because it reveals secrets, but rather because taken together it represents a unique chronicle of the people and their efforts to develop a technology. This then is the focus of the next section on the use of the literature *as a source of data* about what goes on in researcher communities.

Literature as a source of data

A small number of scholars have sought to use the literature as a source of data about emerging technologies and research communities. One pioneering effort is Mullins' 1972 study of molecular biology, in which he took advantage of co-authorship in the literature to understand the

communication network among researchers in the field.[9] In instances as these, the investigators went beyond numbers of publications and instead probed the literature for specific information contained therein: that is, about the people involved and their interaction with each other, or the nature of the work itself.

Journal articles, conference papers, and patents generated by researchers in a given field indeed represent a detailed, self-reported archival record of their efforts to solve the scientific and technical problems confronting them. Indeed, conventions for documenting new knowledge in the literature have resulted in the inclusion of a wealth of secondary information in each document. Once produced, this information can be extremely helpful in measuring a number of structural and behavioral characteristics of a community and how they change over time. It is readily apparent that there is much more information contained in a journal article, or for that reason in a patent, than is usually taken advantage of in a traditional bibliometric analysis.[10] The typical article informs us of the people involved, the topic and nature of the research conducted, where the research was performed, and roughly when it was completed. But an article can contain additional information, either directly or inferred from the basic raw data. For example, as shown in Mullins' study, analysis of co-authorships can provide information on the incidence of collaboration. The acknowledgements in an article can reveal similar information. Typically acknowledgements list

---

[9]N.C. Mullins, "The Development of a Scientific Specialty: The Phage Group and the Origins of Molecular Biology," *Minerva* 10 (1972): 51-82. Since Mullins, others have sought to use co-authorship data to investigate collaboration among researchers, including constructing a map of the interorganizational linkages within a community.

[10]Most journal articles are standardized in format, and can include all or some of the following raw data:

- name of author(s)
- organizational affiliation(s) of author(s)
- geographical location of author(s)
- title of article
- abstract
- keywords
- country of publication
- acknowledgements
- sponsors
- author biographies
- content of article
- references
- journal name
- language
- publication date
- submission date
- special footnotes

researchers, both inside and outside the authors' organizations, who have contributed to the development of the knowledge generated in the paper. Special footnotes can indicate when a researcher changed organizations in the interval between submission and publication of the article. Moreover, longitudinal analysis of articles authored by a given researcher may indicate permanent employment changes.

The literature data are preferably organized into a relational database since this facilitates longitudinal analysis of the data. Software packages for building customized relational databases are easy to obtain and can be turned to the researcher's advantage. Each article, when entered into a relational database can then be analyzed to yield the following structural and behavioral dimensions of the R&D community under study:

- the growth (or decline) in participation of researchers in a community [11]
- analysis of subsets of the community [12]
- fluidity within a community (by studying entry and exit patterns of researchers)[13]
- mobility of researchers between different organizations
- interorganizational information network

---

[11] Obviously, the manipulations that can be carried out with this structural characteristic are numerous. Standard manipulations are: indexing the growth data through the use of a chosen reference base-year (this makes comparison across different R&D communities feasible) -- computation of compound growth rates for a particular field -- aggregating the number of researchers in two (or more)-year time intervals (in order to smooth the data and to account for the fact that many researchers only publish once every two or three years) -- introducing time-lags to account for the fact that a certain period of time (which may vary from field to field) elapses before a piece of research is ready to publish.

[12] A typical subsets is: the core group of researchers, i.e. the researchers who show a steady involvement in the field, who have been present for long times, and who are the promotors of the newly emerging discipline. This can typically be done by looking at those researchers who appear at regular time intervals in the database. Other relevant subsets are: different areas of specialization in a field, geographic clusters of researchers, etc.

[13] As far as entry/exit patterns are concerned, one can recur to quantitative techniques such as the computation of survival rates.

- rate of diffusion of community among different organizations, sectors, and countries[14]
- diffusion of particular ideas or methods and adoption of standardized approaches[15]
- degree of specialization among organizations on particular problems or applications[16]
- technological leadership or followership of different organizations and countries[17]

Using the literature in such a way is particularly appealing. It allows one to map out the field comprehensively and unobtrusively along a large number of structural and behavioral dimensions. The statistics just described are important, though not the only ones in which one might be interested. Managers might for instance choose to track the activities of their main competitors. Questions then become: when did a particular competitor enter the field, how many man-years did he already invest in the new technology, on what particular areas of the new field is he concentrating, with whom is he collaborating etc. The data are publicly available and thus easy to obtain at relatively low cost. Moreover, looking at the literature in this way circumvents many of the problems inherent in citation analysis and simple volume number counting. This can be explained as follows:

---

[14] The diffusion rate can be computed through the definition of a concentration index of researchers in organizations (two readily available indices are the percentage of researchers working in the five largest research groups for any given year, or the computation of a Herfindahl index $H=\Sigma(S_i)^2$ with $i=1 \ldots n$ and S_i the percent of researchers employed by the ith organization). Changes in the concentration index over time are then interpreted as a reflection of the diffusion of the community.

[15] For instance, in the study of neural network technology (see Rappa and Debackere, 1989) the diffusion of the back-propagation algorithm as learning rule in neural networks can be easily traced from its inception in 1986 to the present day be looking how many researchers claim their network model to be based on a back-propagation learning rule.

[16] One obvious way to do this is by tracking the researchers belonging to so-called new ventures who specialize on one particular subset of the global problem set. For example in the neural network field, one can identify research groups solely working on the development of dedicated hardware for neural network implementation, while others may be working on the development of better learning rules etc.

[17] Obviously, this also allows one to see where the centers of gravity of the development of a new technology are situated, and still more interesting, how they may change over time.

- The quality of the data. The conventions of publication and patenting are such that one can be fairly confident in the accuracy and authenticity of the data: that is, among other things, the people and organizations are indeed those reported. Moreover, the peer review process of journals and conferences, and the legal review of patent applications, assure that the research activity itself is of sufficient quality that it is worthy being reported.
- Through the use of objective data in the article or patent, the problems concerning the behavioral motives for citing other research (as present in citation analysis) are alleviated.
- The quality-of-publication problem as present in volume number counting bibliometric applications is avoided since each researcher is only considered once a year, regardless of the number of papers or patents he may have produced in any given year.

3. Limitations of the use of the literature as a source of data on emerging technologies

Although the previous discussion has highlighted some of the benefits of analyzing the information contained in the documented literature in the process of monitoring emerging technologies, a few cautions are warranted before one intends to use this technique. First, some cautions pertaining to the technique in general are discussed. Second, as the analysis of a field can be facilitated through the use of electronic database retrievel systems, the limitations of electronic databases retrieved via abstracting services deserve attention. Third, some additional difficulties specific to the inclusion of patents in a database are highlighted.

General remarks on the use of the literature as a source of data

The only notable difficulties with bibliometric data in general are with identifying and gathering the source documents, and constructing the database to be analyzed. Identifying the boundaries of a field can be a challenging task and usually requires close study of the basic ideas involved. However, since researchers in a field are also confronted with the challenge of locating their respective body of literature, it usually does not take long for an emerging field to develop a

specialized vocabulary of key terms, as well as specialized journals and conferences, which then can be used to locate the majority of the relevant literature. However, one has to admit that the boundaries of a field will always remain fuzzy to a certain extent. There is also a certain time lag between the completion of the research and its publication in a journal or a patent. This makes bibliometric data well suited for historical analyses of the process of the emergence of new technologies, though it is less suited for real-time analysis.

Once the boundaries of the field are sketched, the source documents have to be collected and the raw data have to be entered into the computer database. As is to be expected of any large-scale archival data collection, this is a particularly tedious effort. This is particularly true in the case of a field with a long history or one which is relatively large, perhaps encompassing several thousands of documents.[18] In addition, entering the data into a suitable computer database is labor intensive and subject to typographical and other data entry errors.

The time-consuming burden of identifying and physically collecting documents and creating a database is alleviated by the availability of bibliometric data from database vendors. These companies provide online computer access to many different literature databases. However, while the data collection process is certainly easier, there are a number of trade-offs to be made in using these commercial online databases. These trade-offs are the focus of the next section.

---

[18] For example, a study of compound semiconductors required the identification and analysis of 5,905 journal articles and conference presentations on the subject of this technology published between January 1974 and March 1987. See M. Rappa, "Assessing the Emergence of New Technologies: the Case of Compound Semiconductors," in *Research on the Management of Innovation* A. Van de Ven, H.L. Angle and M.S. Poole (ed.): 439-64 (New York: Harper and Row Publishers, 1989).

The use of electronic databases

The ease with which the source data can be collected is a major advantage of the use of commercial electronic databases. Speeding up the process of data collection has obvious advantages: the number of fields under study can be greatly expanded thus adding to the external validity and generalizability of the frameworks the researcher want to explore by means of this bibliometric approach. Of course, the difficulties previously discussed remain. Some additional ones may even appear. Though, before turning to these, the collection of the source documents once again deserves attention.

The problem of fuzzy disciplinary boundaries does not disappear with the use of electronic databases. As the source documents are collected using a computerized search strategy on a number of keywords (or Boolean combinations of these), one has to be very careful about the body of literature abstracted from the retrieval service. Obvious problems are the collection of "garbage documents", i.e. documents that do have nothing to do with the field studied.[19] Another problem is the collection of articles which have been published in trade or commercial-type journals and hence do not reflect any real R&D activity. Yet another issue is the search strategy itself. Slight changes in the search commands may yield huge differences in the collection of retrieved data documents. Therefore, experimentation with a number of relevant search strategies is desirable before deciding on a particular retrieval strategy. Once one has decided upon the search strategy most appropriate to a particular field, the documents can be retrieved.

Given the possibility of "garbage documents", each selected document has to be studied carefully to see whether it is suitable for the research purpose. Thus, although one can save time in

[19] Obviously, the introduction of Boolean searches can alleviate this problem only to a certain extent. For example, in retrieving a database on polypropylene catalysts, a careful analysis of 3,400 retrieved documents revealed that only 948 of them dealt with catalyst development, although the search strategy was geared towards the specific types of catalysts used in polypropylene manufacturing.

retrieving a huge number of abstracts via electronic means, a number of tedious manipulations are not likely to be solved. Beside the check on the documents themselves, there is the obvious need to standardize the database. Although data retrieval services are steadily developing conventions for documentation (similar as the standard conventions in journals or patents), they still have some way to go. This means that the researcher who intends to use an electronically retrieved database has to do some typographical work in order to standardize the database. By standardization we mean such interventions as:

- designing a unique name to each organization throughout the database (at the moment, a particular organization can can easily appear under three different description, for example: Philips International NV, Nederlandse Philipsbedrijven, Laboratoires Electronique Physique);
- checking the spelling of author names (given our use of the database, this is of tremendous importance. Here one has to recur to the original hard-copy, if necessary. For example is A.V.W. Smith who published in 1987 identical to A.W.V. Smith who published on the same topic a year later?);
- inserting additional standard codes such as *Co.* for an industrial organization or *U* for a university (this enables one to easily group particular subsets of researchers, for example university versus industry researchers. These are of course not the only subsets possible. One might be interested, for instance, in particular sub-fields, countries, applications of the technology etc. It is up to the researcher has to determine in what he is interested and to adapt the database accordingly so as to facilitate database manipulations);
- completing documents with missing information (cfr. infra, when discussing the problem of missing inventor names on patents).

It is clear that, although less tedious than collecting all the relevant source documents

themselves, the electronic database avenue is not without problems itself. Moreover, even if the retrieved database has been cleaned up and standardized, the amount of information stored in each document is still only a fraction of the information contained in the original hard-copy. For instance, the electronic version does not contain the entire contents of the document, but only the abstract. Other significant omissions are: (1) absence of affiliations of secondary authors; (2) absence of submission dates; (3) lack of acknowledgements or special footnotes; and (4) lack of author biographies. Furthermore, at the moment, the documents contained in electronic databases typically extend back only to the 1960s, when many of them were initially created.

The use of patents as a source of data

Patents are a part of the total body of documented information which are increasingly popular in bibliometric analyses of technological activity. Patents can be manipulated in exactly the same manner as journal articles and conference presentations when one wants to use the objective data contained in each document. The previous caveats remain valid. For instance, if one intends to use electronic patent databases (such as Derwent or Inpadoc), the determination of a good search strategy is of first-order importance. Cleaning up and standardizing the obtained documents remains necessary. However, patents in and of themselves appear to have some additional intricacies which hampers their use to some extent.

As the technique proposed here goes beyond volume counting or citation analysis, and tries to use the information present inside the document, the quality and completeness of this information is of crucial importance. Normally, the legal examination of patent applications warrants the minimum level of quality necessary for the proposed methodology. However, the use of a number of (chemical) patent databases has shown us that a lot of patents are stored in electronic databases without inventor names. This absence of inventor names is particularly problematic since within a specific R&D community, the individuals are the focus of the analysis.

Luckily, a combination of different databases allows one to overcome this problem and to complete the missing inventor names in the source database. This, however, is a tedious process. Roughly, one has to identify the patent application number on the documents in the source database which have missing inventor names. Next, these patent application numbers have to be entered into a search procedure on another electronic retrieval system. For example, in a study of two chemical technologies, we derived the source database via Dialog. The missing inventor names were detected and completed using the Derwent and Inpadoc services. Once one has the lacking inventor names, they have to be entered into the original source documents. As can be imagined, this is a process that may keep one busy for a while, in particular when the field is large or when it has a long history.

Another limitation to the use of patents stems from the fact that patents which have been applied for, though never got issued, do not appear in electronic databases. Thus, the researcher may actually only unveil the top of the iceberg in a certain field. A good understanding of the field can give the researcher some insight into whether the problem is serious for the R&D communtiy under study or whether the problem is only marginal. Finally, intercountry differences in the speed of issuing patents and in the time lag between the filing of the patent and its publication (although 18 months appears to be a rather common delay period) require the researcher's attention. A strategy one can use to alleviate this problem consists in taking the priority date on the patent as an indicator of the year in which the inventors were active in the field.[20] Furthermore, given time lags between patent filing and subsequent publication (18 months or more), patent data are not extremely useful for real-time monitoring of R&D communities. But, this is a remark which holds for the published literature as well.

---

[20]This approach is advocated by others as well. See for instance N. Lambert, "Patent Statistics Searching: Practices and Pitfalls," *Proc. of the International Chemical Information Conference* (1989): 141-60.

4. Discussion: application possibilities of the method

Bibliometric data, when analyzed according to the method described, yield a wealth of statistical, empirical information on the growth of a particular R&D community. The obtained information can be used at two levels. This is now briefly discussed.

A first level is perhaps the most obvious one. Although a content-analysis of bibliometric documents has limitations when one wants to obtain real-time information on the technology, a major benefit of this kind of analysis is its ability to make a dynamic quantification of the process of the development of a particular technology. The longitudinal information allows managers or policymakers to see how certain structural and behavioral characteristics in a particular R&D community have evolved over time. And, since we intend the method to be suited to understand the process of the emergence of new technologies, the real-time limitation does not seem fatal. Indeed, it takes decades for a new technology to emerge. Understanding how the community evolves over such a period is obviously more important than knowing what happened the last 6 months or so. Monitoring new technologies, it is our firm belief, is an effort requiring long-term commitment. Thus, the statistics previously described generate a lot of useful, "intelligence" information on a specific field of interest.

However, the second level at which the bibliometric statistics can be useful, is the one that fascinates us most. At the outset of the paper, it was outlined that recently researchers have come to formulate the hypothesis that new terchnologies emerge within the context of R&D communities. In other words, the emergence of a new technology is perhaps not so much an organizational activity as it is a communal activity. While it is true that individual firms play an instrumental role in technological development, the emergence of a new technology is more likely to be the result of a broader intellectual movement, one that embraces certain ideas and techniques in common and extends across a diverse collection of organizations wherever they may be located throughout the

world. Perhaps this is best illustrated by Braun and Macdonald, who in describing the development of the transistor at Bell Laboratories and the advent of semiconductor technology claim: "It is ...unrealistic to see the transistor as the product of three men, or of one laboratory, or of Physics, or even of the forties. Rather, its invention required the contributions of hundreds of scientists, working in many different places, in many different fields over many years."[21] The statistical information obtained from the bibliometric analyses allows one to start looking into the sociological dimensions of R&D communities. The study of R&D communities from this perspective enables one to understand their dynamic (time-dependent) structural and behavioral characteristics. Once this obejective is reached, we can start thinking of how these dynamics are linked to the rate of technological progress. Answering the question about a possible relationship between structure and conduct in an R&D community and the subsequent rate of progress of the technology is a challenging task. The bibliometric method proves to be an invaluable help.

The consequences of a well-articulated Structure-Conduct-Performance framework for R&D communities are far-reaching. It may alter a firm's approach towards corporate R&D, it may influence national science policies, and most important of all, it may offer us a better insight into the processes going on within the black box which technology development still too a large extent is today.

=§§§=

[21] E. Braun and S. Macdonald, *Revolution in Miniature: the History and Impact of Semiconductor Electronics*, 2nd Edition (Cambridge, UK: Cambridge University Press, 1982)

7659 .75

Date Due

NOV 1 1 1990
MAR 21 1991
JUL 28 1991
JAN 02 1994

Lib-26-67

MIT LIBRARIES DUPL 1

3 9080 00658182 8